Android authorship attribution through analysis of String *n-grams*

Vaibhavi Kalgutkar, Natalia Stakhanova, Paul Cook
Faculty of Computer Science, University of New Brunswick

ISCX
Information Security
Centre *of* Excellence

UNB

Problem Statement

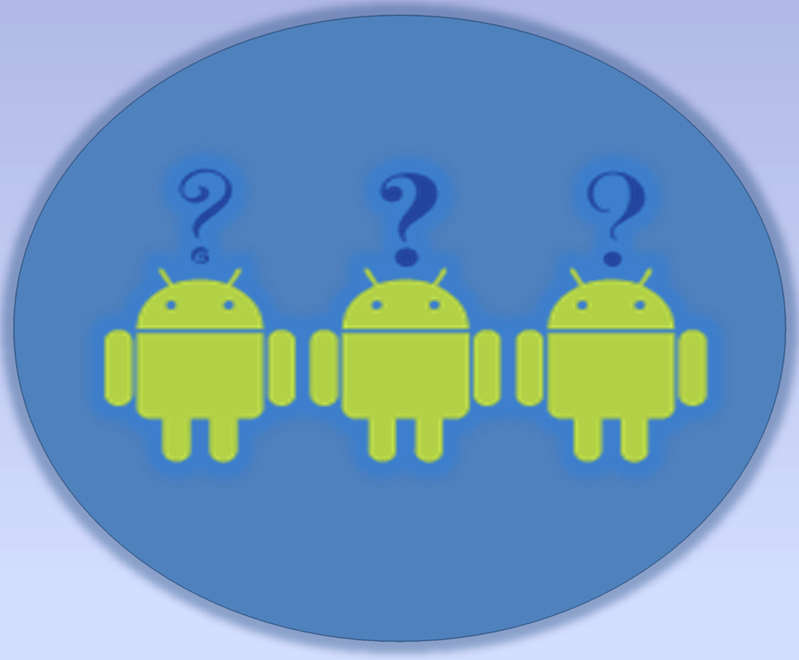
- Mobile device market, especially Android is expanding rapidly
- Increasing number of malicious apps due to openness of google play store
- To mitigate the risk of malicious apps, it is extremely important to understand the motive of the attacker.
- Authorship attribution can help to answer such issues and to minimize the risk of exposure to malicious apps.

Methodology

- A machine-learning based approach
- 3-gram word counts are considered
- Three kinds of strings are analyzed namely referenced, unreferenced and application specific strings
- Impact of these strings on the task of classifying android apps is studied

Experimental Setup

- ✓ Dataset : 1684 apps by 43 different authors
- ✓ Linear SVM classifier
- ✓ 5 times 5-fold cross validation



String Analysis

- Our research focuses on the different text component found in the APK files. We have explored the following string components of APK.
- **Referenced strings present in DEX file:**
  - ✓ Referenced by one of the identifier sections of DEX file
  - ✓ Part of functional app code
- **Unreferenced strings present in DEX file:**
  - ✓ Present in the data section of DEX file and only referenced by string offset list
  - ✓ Carry hidden or interesting textual information
- **Strings extracted from strings.xml** :
  - ✓ Referenced from the application or from other resource files in APK
  - ✓ Application specific strings defined by the author

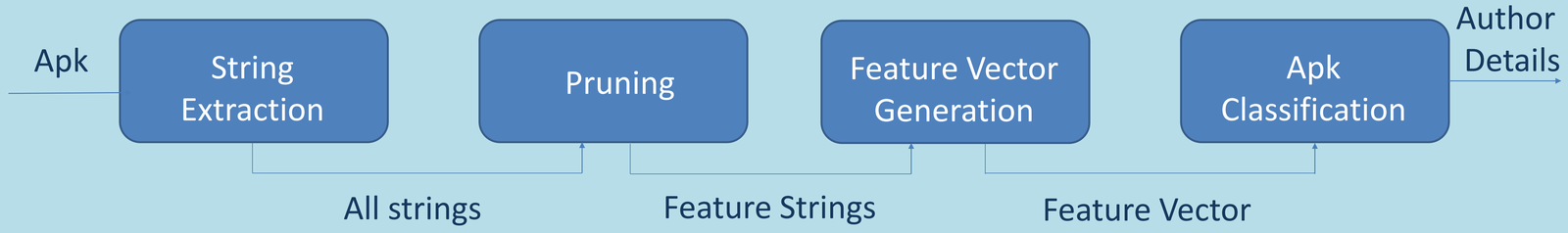


Fig 2. The Flow of String Analysis

Conclusion

We have presented a solution to identify the author of an android app through the use of text strings extracted from the Android Executables file. The proposed system using a Linear SVM with line bounded word level 3-grams was able to identify the authors with an accuracy of 95.52%

Experimental Results

String Type	Average Accuracy	Macro Average Precision	Macro Average Recall	Macro Average F1
Application specific	0.9337	0.9426	0.9186	0.9201
Unreferenced	0.9552	0.9467	0.9392	0.9381
Application specific (tf-idf)	0.9302	0.9377	0.9147	0.9155
Unreferenced (tf-idf)	0.9547	0.9461	0.9384	0.9374
Referenced + unreferenced	0.9616	0.9564	0.9477	0.9477